Title: Manually-controlled Instrumented Spasticity Assessments: a systematic review of psychometric properties

Authors:
Lynn Bar-On PT[a,b], Erwin Aertbeliën Ing PhD[c], Guy Molenaers MD PhD[a,d,e] , Bernard Dan MD PhD[f], Kaat Desloovere PhD[a,b]

a. Clinical Motion Analysis Laboratory, University Hospital Leuven, Belgium
b. KU Leuven Department of Rehabilitation Sciences, Leuven, Belgium
c. KU Leuven Department of Mechanical Engineering, Leuven, Belgium
d. KU Leuven Department of Development and Regeneration, Leuven, Belgium
e. Department of Orthopedics, University Hospital Leuven, Belgium
f. Department of Neurology, University Children's Hospital Queen Fabiola, Université Libre de Bruxelles, Brussels, Belgium

Acknowledgements:
Lynn Bar-On is supported by a grant from the Doctoral Scholarships Committee for International Collaboration with non EER-countries (DBOF) of the University of Leuven, Belgium. This work was further supported by a grant from for Applied Biomedical Research from the Flemish agency for Innovation by Science and technology (IWT-TBM: grant number 060799).

Disclosure Statement:
We certify that no party has a direct interest in the results of the research supporting this article. We certify that we have no affiliations with or financial involvement (e.g. employment, consultancies, honoraria, stock ownership or options, expert testimony, grants and patents received or pending, royalties) with an organization or entity with a financial interest in, or financial conflict with, the subject matter or materials discussed in the manuscript.

Word count: 4675

Corresponding author:
Lynn Bar-On
Clinical Motion Analysis Laboratory
University Hospital Leuven
Weligerveld 1
3212 Pellenberg
Belgium
Tel. +32 16341295 or +32 16338024
lynn.1.bar-on@uzleuven.be

Running title: Manually-controlled instrumented spasticity assessments

Abstract

Aim: The first aim of this study was to systematically review and critically assess manually-controlled, instrumented spasticity assessment methods that combine multidimensional signals. The second aim was to extract a set of quantified parameters that are psychometrically sound to assess spasticity in a clinical setting.

Method: Electronic databases were searched to identify studies that assessed spasticity by simultaneously collecting electrophysiological and biomechanical signals during manually-controlled passive muscle stretches. Two independent reviewers critically assessed the methodological quality of the psychometric properties of included studies using the COSMIN guidelines.

Results: Fifteen studies with instrumented spasticity assessments met all inclusion criteria. Parameters which integrated electrophysiological signals with joint movement characteristics were best able to quantify spasticity. There were conflicting results regarding biomechanical-based parameters that quantify the resistance to passive stretch. Few methods have been assessed for all psychometric properties. In particular, more information on absolute reliability and responsiveness for more muscles is needed.

Interpretation: Further research is required to determine the correct parameters for quantifying spasticity based on integration of signals and especially focusing on decomposing the neural from non-neural contributes to increased joint torque. These parameters should undergo more rigorous exploration to establish their psychometric properties for use in a clinical environment.

INTRODUCTION

Excessive and uncontrolled spasticity causes pain, limits functional recovery and is thought to cause secondary complications such as contractures and bone deformities.[1] It appears in conditions with upper motor neuron (UMN) syndrome and is the most common neurological feature in persons with cerebral palsy (CP). Despite the impact of spasticity and the many therapeutic paradigms aimed at treating it, there are few clinically-suitable, reliable methods for its assessment. One reason for the lack of consensus on the assessment method originates from the absence of a commonly accepted definition for spasticity.[2]

In 1954, Tardieu and colleagues described the phenomenon of a 'spastic catch' as "a sudden reactive resistance to a fast passive stretch of a spastic muscle".[3] In 1980, Lance was the first to define spasticity as "a velocity-dependent increase in tonic stretch reflexes with exaggerated tendon jerks, resulting from hyper-excitability".[4] Although Lance's definition is most commonly cited, in routine clinical practice, it is nearly impossible to distinguish this definition of spasticity from other positive symptoms of the UMN syndrome. For example, other reflex mechanisms (e.g. cutaneous or nociceptive) could also contribute to increased muscle activation and are difficult to distinguish from the proprioceptive reflex mechanisms described by Lance.[5]

Sanger et al. defined spasticity as "resistance to an externally imposed movement that increases with increasing speed of stretch or rises rapidly above a threshold speed or joint angle".[6] However, here too distinguishing the resistance caused by pathological muscle activation due to a hyperactive stretch reflex from the increased resistance due to passive stiffness is clinically very challenging. Non-neural muscle and tendon alterations also contribute to reactive resistance, especially in persons with UMN syndrome.[7] Changes of the viscoelastic properties of these structures will determine both the stiffness and the velocity-dependence of a movement. Thus, it appears that 'observed' spasticity encompasses multiple phenomena and is not a single pathophysiological entity. In line with this finding, the SPASM consortium introduced a broader definition for spasticity: "a disordered sensori-motor control, resulting from UMN lesions, presenting as intermittent or sustained involuntary activation of muscles''.[5]

The complexity of distinguishing spasticity from other positive symptoms highlights the challenges in developing suitable measurement systems. Firstly, a distinction has to be made between measurements that assess spasticity in a relaxed muscle or during activity. In most clinical settings, spasticity is measured using subjective, easy-to-use ordinal scales that assess the level of resistance felt by the examiner during a passive muscle stretch. Examples of such scales include the Modified Ashworth Scale (MAS)[8] and the Modified Tardieu Scale (MTS).[9] The MTS is considered more valid for the assessment of spasticity as defined by Lance as the resistance is compared during stretches at different velocities. However, lack of standardization of stretch velocity and the subjective nature of both scales has resulted in poor inter-rater reliability[10,11]and, for the MTS, inaccuracies in determining the correct catch angle.[12] In light of the above mentioned difficulty of isolating spasticity, these tests also greatly oversimplify the phenomena. It is therefore not surprising that many studies have shown poor correlations between the clinical measures (MAS, MTS)

and objective indicators of pathologically increased muscle activity during passive stretch.[7,13–15] For example, some subjects, who have been found to have spasticity during a clinical examination as indicated by increased resistance to passive stretch, lacked any signs of hyperactive H-reflexes.[16] In these cases, increased resistance to passive stretch may have been due to non-neural causes.

Therefore, it is now acknowledged that quantified, instrumented methods should be used to provide a more accurate and valid evaluation of spasticity.[17] In 2005, the state of the art on spasticity assessment was thoroughly summarized by the SPASM consortium into three review articles.[18–20] These reviews identified and categorized a large number of non-invasive, instrumented applications for quantitative spasticity assessment into *biomechanical* and *neurophysiological* methods[18,19], and concluded that both methods are complementary and should be used simultaneously to sufficiently differentiate between neural and non-neural causes of increased resistance.[20] *Biomechanical devices* record joint-angular characteristics and/or resistance around a joint during passive stretching.[18] They include for example motor-driven or hand-held dynamometers. *Neurophysiological methods* measure muscle activity using, for example, electromyography (EMG) during passive movement or nerve stimulation.[19] Furthermore, the consortium stressed that collecting experimental data in a highly technical and controlled environment would greatly improve the modeling of the complex pathophysiology. However, combining these recommendations in view of a clinical application requires some compromise. A suitable method should on the one hand be more valid and reliable than the current clinical tests; and on the other hand, remain clinically feasible in different patient populations, including children. For example, whilst some motor-driven, isokinetic devices that measure limb resistance to passive movement have great reliability because the limb is moved at a controlled velocity,[21–24] these are bulky and often difficult to apply to children in high-velocity stretches.[20] In addition, a stretch reflex may be more easily elicited by a transient acceleration which is robotically more difficult to apply.[25] A manually-controlled displacement method offers a clinically-applicable alternative.[26–28] However, to ensure accuracy, manually-controlled displacement methods must follow standardized protocols and the psychometric properties need to be defined before they can be used in clinical practice.[20] A recent review of spasticity assessments for children and adolescents with CP highlighted insufficient psychometric soundness of spasticity evaluation tools.[29] However, this review did not emphasize the need to integrate biomechanical and electrophysiological signals, as is recommended for valid spasticity assessment.[20] Therefore, their conclusion that electrophysiological methods to assess spasticity demonstrate the most promising results in terms of reliability and discriminate validity may have been misleading.

The aim of the current study was two-fold. First, we wanted to systematically and critically assess clinically-applicable spasticity measurement methods that adhere to the recommendations of the SPASM consortium.[20] Following these recommendations, any developed spasticity measurement method should (1) be able to make measurements at variable velocities of displacement; (2) incorporate simultaneous recording of EMG and torque; and (3) include a clearly defined protocol. To ensure a similar conceptualization of spasticity across reviewed articles (i.e. the definition of spasticity as offered by Lance[4]), only measurements during

passive conditions were to be included. Secondly, we aimed to extract a set of quantitative parameters to measure spasticity based on the reviewed articles.

METHODS

Search Strategy

A single reviewer (LB) performed a web-based search for relevant literature using the following electronic databases: Science Direct (www.sciencedirect.com), MEDLINE (PubMed) and Embase (www.embase.com). Only full-paper articles published in English in peer-reviewed journals, performed on human subjects, were included. Keywords included ('All fields' and MeSH): (1) spasticity; (2) tone; (3) cerebral palsy; (4) stroke; (5) spinal cord injury; (6) upper motor neuron; (7) measure; (8) evaluation; and (9) assessment. The following word combinations were implemented: 1 or 2; AND 3 or 4 or 5 or 6; AND 7 or 8 or 9.

Study selection

Two reviewers independently selected the studies for inclusion in the review. First, titles and abstracts were screened for eligibility. Second, the full text of potentially relevant papers was read to ascertain whether the study met all selection criteria, i.e. the article had to describe a method to quantitatively assess spasticity by recording both biomechanical and electrophysiological signals during manually-applied passive muscle stretches. Studies were excluded in case the method (1) only assessed spasticity based on subjective measurements, including Ashworth- and Tardieu-like scales[11]; (2) only applied a motor-controlled device or a pendulum-like test[30] to stretch the muscle; (3) was limited to collecting either biomechanical or electrophysiological signals; (4) applied a passive stretch at only one velocity; or (5) assessed spasticity during function or active movements. Use of the tendon- and Hoffmann reflexes as a means to assess spasticity has been extensively studied,[19] however their clinical applicability and relevance is limited. Therefore, also studies applying excitation of these reflexes or electro stimulation as a neurophysiological means to assess spasticity were excluded from the current review. Finally, in those cases where more than one article was published by the same research group with the same methodology, the most recent publication was selected for review unless older articles investigated different psychometric properties. The bibliographic details of excluded studies were listed and reasons for exclusion noted. Any discrepancies regarding final selection were resolved by consensus and, if necessary, by consulting a third reviewer.

Data extraction and quality assessment

Selected studies were read by two independent reviewers (LB and KD) to extract information on study populations, methodology, study design, outcome parameters, results, and conclusions. Both reviewers independently evaluated the quality of the psychometric properties of the described method using the COSMIN checklist.[31] The COSMIN checklist offers a common terminology and definitions of psychometric properties and consists of 12 domains.[32] For each study included in the current review, only those domains relevant to the investigated psychometric properties were checked. The relevance of each domain and the interpretation with respect to spasticity measurements was discussed prior to commencing. Six domains were considered relevant (Table 1A in SuppInfo1): two were used to determine whether a

study met the methodological quality on reliability and measurement error; two assessed the methods' content and construct validity (including hypothesis testing); one assessed the responsiveness of the method; and finally, one determined the interpretability. Generalizability was determined for each of the previous domains. The following domains from the COSMIN checklist were not considered relevant for spasticity assessment: Item Response Theory (IRT), internal consistency, structural validity, cross-cultural validity and criterion validity. Reasons for not assessing these properties are described in Table 1B in SuppInfo1.

Each of the six domains (and generalizability) were rated by both assessors independently on a 4-point scale according to the COSMIN guidelines.[33] 'Excellent' quality was assigned if all relevant COSMIN items within a domain were scored as adequate. 'Good' quality was assigned to those studies that lacked some aspects, though it could still be assumed that the items were acceptable. 'Fair' quality was assigned if the measurement property was underrepresented, explored in a moderate sample size or when there were other minor flaws in the design or statistical analyses. 'Poor' quality was assigned if there were major flaws in the design or statistical analyses. Finally, in each article, the statistical findings per domain were rated according to quality criteria provided by Terwee et al. (2007) as positive, indeterminate, negative, or no information available (Table 1A in SuppInfo1).[34] Per domain, all items, resulting scores and statistical ratings were then discussed by the reviewers and any discrepancies resolved by consensus.

RESULTS

A flow chart of the selection process can be viewed in Figure 1. After filtering the databases on keywords and screening titles and abstracts, 158 potential full-text articles were found. Further examination of these full-text articles revealed that 33 papers did not apply an objective measurement method, 39 used a robot to displace the limb, 27 applied electrostimulation, and 38 articles measured either a biomechanical or an electrophysiological signal in isolation. One article measured both signals, but did not use the biomechanical parameters as a means to quantify spasticity. One article was excluded as the limb was only displaced at one velocity. Finally, three articles were excluded as their methodology was reported in more recent versions by the same research groups. Therefore, 15 studies were identified as meeting all the inclusion criteria. The data extracted from these are summarized in Tables 2-4 and in SuppInfo 2 and 3. A list of excluded full-text articles can be found in SuppInfo 4.

Study populations and muscles tested

Information on subjects, instrumentation and protocol details are summarized in Table 1. Seven of the 15 articles studied spasticity in adults post-stroke.[35–41] Two articles included persons with spinal cord injury,[42,43] and four reported on children with CP.[27,44–46] One study included adults post-stroke and adults and children with CP in the subject group.[7] One article included adults post-stroke, spinal cord injury and adults with CP.[47] Eight studies additionally included a healthy control group.[27,35,39–41,45–47] Six articles studied spasticity in upper limb muscles[13,27,35,37–39], eight in lower limb muscles,[40–47] and one in both upper and lower limb muscles.[7]

Instruments and protocols

Angular position/velocity was recorded in most studies using calibrated potentiometers or electrogoniometers[7,13,27,35,37,39–44,47], in two studies using inertial sensors containing an accelerometer and a gyroscope[45,46], and in one study, a velocity sensor was used.[38] Forces and/or torques exerted at the joint when manually displacing the segment during passive stretch, were measured with different devices. Most often, force measurements were carried out using single or multiple-axes force transducers[7,13,35,37,39–41,43–47] or differential pressure sensors.[38] Forces were then recomputed to torques based on measurements[39–41,44–47] or estimations[38,43] of moment arms. Three studies directly measured torque near the joints[42] in order to account for the torques applied by the examiner on the handle of the sensor.[45,46] All studies used surface EMG (sEMG) to record agonist muscle activity and eight studies additionally measured the antagonist muscle activity.

All studies assessed spasticity during passive ramp stretches of the spastic agonist muscles, except for three studies that analyzed passive sinusoidal movements[35,38,39], and two studies that did both.[40,41] Stretches were performed either at two velocities (slow and fast)[7,13,37,39,41,43,45,46]; at three velocities[35,44]; or at four or more velocities.[27,38,40,42,47] Stretch velocities ranged from 2-720°/s. One study did not report the applied stretch velocity.[44] Within each velocity, stretch repetitions were applied at zero to one minute intervals.

In addition to instrumented spasticity tests, 12 of the 15 studies assessed spasticity with the (M)AS[7,13,27,35,37,38,41,42,44–46] and two studies additionally used the (M)TS.[44,46] Three studies in adults post-stroke also examined the relation between spasticity indicators and upper limb function.[35,37,39]

Study design and data analysis

While most authors failed to mention how spasticity was defined in their study, the majority followed the reasoning that velocity-dependent hyperactivity of the stretch reflex causes a pathological augmentation in muscle activity.[4] Slow stretching was performed at a velocity below the threshold of stretch reflex activity, whereby it was hypothesized that non-neural elastoviscous muscle properties accounted for any increased force or torque measured over the range of motion (ROM). During a high-velocity passive stretch, activation of the muscle additionally influenced any increase in torque. The amount of gain in muscle activity, its timing and the amount of torque produced at different stretch velocities constituted some of the possible quantifiable measures of spasticity. A summary of the main outcome parameters developed by each study to quantify spasticity can be found in Table 2. In Table 2, a distinction is made between parameters that mostly reflect either angular position/velocity, forces and/or torques, or muscle activity. The velocity at which each parameter was examined is also specified. However, most studies combined different signals and velocities to develop their outcome parameters.

For the angular position/velocity parameters, all, but two[40,43], studies measured the available ROM during a passive stretch performed at a velocity below the threshold of stretch reflex activity. Therefore, any decreased ROM or catch angle[27,46] during a higher velocity stretch was presumed to be caused by increased muscle activity. Often referred to as either resistance[37,40,41,47] or stiffness[27], the slope of the torque-angle curve was the most common measure of increased torque. This parameter was calculated over the entire ROM[13], or over a section of the ROM[35,37,39–42,47] and

compared between velocities[13,27,35,37,40,47] or between positions.[42] Four studies examined the torque value at a specific joint angle at different velocities.[27,43,45,46], Four studies additionally examined the integral of the torque-position graphs to quantify the amount of work needed to stretch the examined muscle[27,45–47] and one study calculated the integral of the torque-time graph.[7] When stretches were performed against the force of gravity and the mass of the displaced segment was not negligible[7,27,38,42,44–47], five studies subtracted the effect of inertia from the resulting measured torque.[27,38,45–47]

Nine of the 15 articles quantified sEMG amplitude by calculating the average root mean square of the sEMG signal (RMS-EMG) over a particular interval,[7,13,35,37,39,42,44–46], two by examining the gain in RMS-EMG over the ROM,[40,41] and one by calculating the maximum value of the RMS-EMG.[47] Similarly to the biomechanical parameters, average RMS-EMG was often calculated over a specific portion of the ROM and compared between velocities. Two articles normalizing the RMS-EMG amplitude value to maximum isometric voluntary contraction.[44,45] Three articles recorded and analyzed either the angle or the velocity at EMG onset.[27,38,42] Two articles identified different types of spasticity based on sEMG parameters.[37,47]

Psychometric properties

Reliability

The COSMIN scores of those studies examining reliability can be found in Table 3. For an extended version of this table also containing the methodological and statistical results and scores the readers are referred to SuppInfo2. Six studies[27,35,39,42,43,45] explored the intra-rater reliability of some, or all, outcome parameters from the instrumented tests and two studies[13,46] referred to previously collected reliability results. Of these eight studies, only four examined the reliability of electrophysiological parameters in addition to biomechanical parameters in patient populations[35,39,45,46] and two studies additionally assessed inter-rater reliability.[35,39] The methodological quality of studies ranged from poor to good as study samples tended to be small or the interval between repeated measurements was inappropriate. Reliability results were generally better among persons with disabilities than among control groups and biomechanical parameters tended to have higher relative reliability than electrophysiological parameters.[35,39] Turk et al. reported on the measurement error of the parameters in their study, which ranged from 40-77% of the mean values of those parameters in their subject sample.[39] Several parameters from the studies by Bar-On et al. were found to have an absolute measurement error small enough to distinguish between groups[45] and detect change due to treatment.[46] The minimally important change (MIC) was not identified in any study.

Validity

The COSMIN scores on the validity of the different studies are summarized in Table 4. The methodological quality of the included studies ranged from poor to excellent with the main weaknesses being uncertainty of statistical strength and limited analyses mainly for content validity. Reasons for score allocation per domain together with methodological and statistical scores can be found in SuppInfo3.

Content validity

Content validity was evaluated by a comparison of biomechanical to electrophysiological parameters,[7,13,37,40,47] or by a comparison of parameters between

stretch velocities.[7,13,27,37,40–42,45–47] Pandyan et al.[13] and Fleuren et al.[7] reported conflicting results regarding the correlation between RMS-EMG and the slope of the torque-angle curve in spastic elbow flexors.[7,13] On the hand, in the soleus of subjects post-stroke, higher torque values were associated with hyperactive stretch reflexes[40] and the gain in EMG accounted for 27% of the variance in the measured torque.[41] Associations between patterns of muscle activity and the biomechanical parameters during high velocity passive stretches could not be demonstrated in the wrist[37] or the knee flexors.[47] On the other hand, electrophysiological[13,27,37,40,42,45,46] and biomechanical[7,37,38,45,46] parameters often changed with increasing stretch velocity. Two studies reported no increase in the slope of the torque-angle curve between velocities.[13,40]

Construct validity and hypothesis testing

Evidence of the constructs or hypotheses were tested in 12 studies by either comparing persons with disabilities to a control group[27,35,39–41,45,47], by comparison to a clinical spasticity test[7,13,27,35,37,41,44–46] or by comparison to a motor-driven test.[35,43] In those studies comparing persons with disabilities to controls, average RMS-EMG parameters were always able to distinguish between groups.[35,39,46,47] In contrast, only in four studies, and only in some muscles, were biomechanical parameters able to distinguish persons with disabilities from controls.[27,40,45,47] Conflicting results were found when outcome parameters were related to the scores of clinical spasticity tests. Two studies reported good, significant correlations (r=0.64) between RMS-EMG and MAS-scores for some muscles[7,35] while others reported low associations (r=0.06[13], k=0.09[44]). RMS-EMG parameters were significantly higher in hamstring muscles of children with CP with high MAS scores (2-3) than those with low MAS scores (1-1+), but this was not the case for the gastrocnemius.[45] Similarly for the (M)TS, conflicting results were found for the calf muscles of children with CP with one study reporting good agreement (k=-0.48) between the angle of response as measured by the TS and RMS-EMG[44] and another, only poor to fair (r=0.2) correlations.[45] In five studies, ROM and biomechanical parameters were strongly correlated to MAS-scores[7,27,35,41,45] and in one study to the TS.[44] However, Malhotra et al.[37] found that their biomechanical parameters did not increase with increasing MAS-scores. Bar-On et al. found that the instrumented assessment identified significantly more responders to treatment with Botulinum Toxin-A injections in the hamstrings than the MAS, but not more than the MTS. However, a combination of several baseline parameters from the instrumented test could better predict the effect of treatment than the baseline MTS alone.[46] Parameters from a manual device were compared to those from a motor-driven device and showed very good correlations (r=0.86-0.94).[35] On the other hand, Lamontagne et al. detected fewer subjects with hyperactive stretch reflexes using the motor-driven system than with the hand-held device although, in this study, stretch velocities were not comparable.[43]

Responsiveness and interpretability

Responsiveness to anti-spasticity medication was evaluated by only two studies. However, the conclusions of one study were weakened as the methodology did not fulfill all criteria for high quality.[38] No study provided minimally important change values. In three studies,[39,45,46] the smallest detectible change (SDC) values could be calculated from the reported absolute measurement errors. Bar-On et al. identified EMG and torque-related parameters that, relative to the SDC, decreased post-

treatment.[46] No study investigated all aspects of content validity, construct validity and responsiveness as relevant to spasticity measurement.

DISCUSSION

The goal of this systematic review was to identify instrumented spasticity assessment methods that could be used as viable alternatives to the commonly-applied clinical evaluations such as the MAS. Fifteen instrumented spasticity assessment methods developed following the recommendations by the SPASM consortium[20] were identified. These methods are manually-controlled, ensuring their ability to be translated to clinical settings, and measure both electrophysiological and biomechanical signals.

In comparison to previous reviews[17–20,29,30], the current paper covered a narrower scope of spasticity assessments by reporting on the measurement of passive-state spasticity only. This focus ensured that the concept of spasticity was similarly defined in all of the included studies, namely the definition of spasticity as offered by Lance.[4] A wider definition of spasticity includes spasticity as manifested during active conditions.[5] The exact pathophysiology of spasticity during active motion remains debatable,[48] and consequently, the literature related to its impact on function, divided.[40,49] While in the passive state, enhanced muscle activity is primarily pathological, in the active state, it is more difficult to discern reflex-mediated activity from voluntary activation. In persons with an UMN syndrome, activation is also influenced by other phenomena such as sensory-motor control problems and weakness. It is therefore speculative whether one can apply a theory developed for measurement of a phenomenon in the passive state to the complex activation occurring during activity.[50] While it is acknowledged that spasticity affects activity, we believe that accurate assessment methods need first to be developed for passive and active situations separately in order to decompose the multifactorial phenomenon.

Overall, findings of the current review show that manually-controlled instrumented spasticity assessments that are clinically-applicable are available. Those developed for assessing spasticity in the hamstrings in children with spastic CP, have, so far, undergone the most rigorous clinical assessments.[45,46] However, no developed method has been sufficiently assessed on all the required psychometric properties. Several UMN syndromes were assessed in the included studies showing that spasticity can be quantified in a variety of different pathologies. However, most literature on this subject has been carried out in adults post-stroke and the number of muscles investigated remain limited. This indicates that instrumented spasticity assessment in other areas still requires much development. Similar to the findings of Flamand et al.[29], only six studies were identified studying spasticity in children with CP with information on absolute reliability and responsiveness limited to work by only one research group.[45,46]

Most of the reliability findings were limited to biomechanical parameters with only four studies including a reliability analysis of RMS-EMG parameters among persons with disabilities.[35,39,45,46] Since no or little electrophysiological response is expected when passively stretching healthy muscles, it was not surprising that relative reliability in

control subjects was poor. However, also among patient populations, the electrophysiological response was occasionally found to be variable and unstable.[39] To reduce the variability inherent to RMS-EMG and to be able to compare between subjects, signals can be normalized to a maximum voluntary contraction as was done in two of the reviewed studies.[44,45] However this normalization technique in persons with co-contraction and weakness is debatable.[51] EMG can also be normalized to an M-wave during a supramaximal stimulation.[52] However, more studies are required to assess the clinical applicability of such a method. As an essential start, better protocol standardization is required to reduce the variability of RMS-EMG parameters. On the other hand, the variability in response may also be a true phenomenon of spasticity. More reliability studies are required to investigate this.

Quantification of the measurement error of an instrument is also an important part of reliability and responsiveness analyses. Calculation of the SEM was carried out in three of the reviewed studies.[39,45,46] This permits the calculation of the SDC which is the value of the amount of change that falls outside the measurement error of an instrument.[53] This is essential for a methods application as an evaluative measure in intervention studies and without it, clinical practice is limited. In Bar-On et al., three parameters were identified that, on average, decreased more than the SDC post-treatment with Botulinum toxin-A. In addition, the baseline values of these parameters were able to predict the response post-treatment.[46] The MIC refers to the change which is considered to be minimally important by patients and clinicians.[53] The MIC differs from the SDC as it cannot be statistically determined. Instead, it requires large, in-depth intervention studies often in combination with clinical consensus. Such methodology was not applied in any of the reviewed studies which resulted in limited scores on the interpretability item of the COSMIN checklist.

To be clinically applicable, an assessment also needs to be compact and easy to administer. Although clinical feasibility and utility were not systematically assessed in the current review, the choice to only include manually-controlled assessment methods partially covered this issue. Especially in children, and particularly during high-velocity displacements, a motor-driven device may prevent the subject from being sufficiently relaxed. Manual assessments on the other hand are better tolerated, allow the examiner to have more control over the state of the subject and are transportable. In the study of Malhotra et al.[37], for example, the assessments were performed at the patient's bedside.

The compromise between accessibility and accuracy is also challenged by the necessity to record and synchronize both electrophysiological and biomechanical signals. Fortunately, technological advancements have improved the accuracy, synchronization capabilities and portability of equipment. For example, wireless inertial measurement units are reliable and valid in motion analysis[12] and are recently, being combined with EMG sensor technology.

Recording kinematic data is essential for comprehensive spasticity assessment. First, it ensures the consistency of stretch performance and allows for interpretation of data in accordance to the velocity of stretch. Secondly, with advances in musculoskeletal modeling, kinematic data can be used to calculate muscle lengths and lengthening velocities,[54] essential for spasticity interpretation. While all of the reviewed methods acknowledged the need to assess spasticity at various muscle lengthening velocities,

only eight studies integrated the information from EMG and torque with velocity. Even fewer explored both signals relative to joint position or muscle length.[37,42] Evaluating EMG response to both increasing muscle length and lengthening velocity allows identification of stretch reflex thresholds (SRTs) which in persons with an UMN syndrome, have been found to be reduced.[55] Studies in adults suggest that decreased SRTs may be related to spasticity severity,[56] type of motor deficit,[55] and risk of developing contractures.[37] Investigating both the dynamic and static SRTs in elbow flexors, Jobin and Levin found more velocity-dependence of the SRTs in children with CP compared to adults with stroke.[57] Van der Salm et al.[42] highlighted position-dependent activation in persons with SCI in which the joint angle, rather than the angular velocity, was the trigger of the neurological response. These findings were supported by two more studies that identified either position or velocity-dependent muscle activation patterns among different subjects.[37,47] Chen et al. reported an increase of the dynamic SRT post BTX-A treatment.[38] However, identification of SRTs is highly dependent on the performance of controlled, yet variable stretch velocities which may be more difficult to achieve with manual stretches.[58] Nevertheless, as protocols become more standardized, the reliability of acquiring these parameters with a manual test is worth further investigation.

Several studies were able to show that measuring average RMS-EMG, either over the full ROM, over a specific interval or as a function of velocity, distinguished between persons with disabilities and controls.[19,39,45,47] On the other hand, only three studies showed that some of the developed biomechanical parameters, namely the slope of the torque-velocity curve[27] and the integral of the torque-angle curve[45,47] were higher in persons with disabilities than in controls. Results on content validity showed only moderate correlations between torque-angle curves and RMS-EMG.[13] Chen et al.[38] found that the velocity-dependent viscous component calculated from the torque-velocity curve during a sinusoidal motion was sensitive to treatment with BTX-A. Interpreting these results together, it is possible that a parameter based on torque and velocity best corroborates the velocity-dependent nature of spasticity while the slope of the torque-angle curve is better used as a measure of non-neural related stiffness. The lack of agreement on which parameter best quantifies the biomechanical effect of spasticity may be solved by better differentiation of the neural and non-neural components of increased torque. Models that differentiate into components such as reflex-mediated torque, stiffness and viscosity have mostly been validated on data collected in research settings using motor-driven devices.[24,59] Proponents of motor-driven spasticity assessment devices, argue that by allowing a robot to control the displacement, the limb dynamics of the experimenter can be avoided allowing for accurate modeling of the persons passive state. Nevertheless, as was partially shown by two of the included studies, by improving the performance standardization of manual-tests, a distinction can be made between an increase in torque which is aggregated by muscle activity or an increase in torque of non-neural origin, e.g. contracture.[37,47] Future work should focus on validating the different components and checking their responsiveness to treatment.

Although comparison of an instrumented test to a clinical comparator was indicated in the current review as comprising a part of construct validity, multiple studies have shown the inadequacy of clinical tests such as the (M)AS and (M)TS in assessing spasticity.[10,11,45,60,61] Therefore, it was not surprising that in general, the articles reviewed reported poor correlations between the electrophysiological findings of the

instrumented tests and the scores of the (M)AS and (M)TS. This finding confirms the inadequacy of the clinical tests rather than highlighting the construct validity of the instrumented alternatives. The (M)AS and (M)TS may be useful for diagnostic and broad screening purposes for distinguishing spastic from healthy muscles and for categorizing muscles into broad severity categories.[45,62] However, for a comprehensive picture of the problem and better differentiation of mid-range severities, the clinical exams should be supported by more rigorous, instrumented assessments, especially for persons undergoing treatment.[46]

In conclusion, the search for a clinically-applicable, instrumented spasticity assessment is still ongoing as the translational capabilities from research to clinic are unnecessarily lagging behind. Some promising developments of instrumented spasticity assessments that integrate signals have been found. However, more consensus is required on the optimal parameters that quantify spasticity, provide insight on its nature and differentiate it from non-neural related increases in torque. Parameters based on RMS-EMG fulfill aspects of validity in adults post-stroke[13,37] and in children with CP.[27,45] However, the inter-rater reliability of these parameters remains unexplored and responsiveness studies should be expanded to more muscles and different patient populations. Most importantly, for a parameter based on RMS-EMG to be used as a quantifiable measure of spasticity, methods should aim at standardizing their tests to ensure adequate reproducibility. Few developed torque-related parameters possess convincing content or construct validity to be used as clinical measures of spasticity. However, by improving the joint torque models and differentiating the components of increased torque, this could be achieved. Simple, but accurate applications of an instrumented spasticity assessment will greatly advance clinical practice in terms of treatment planning and outcome evaluation. In parallel, collection of instrumented data will help define and classify different aspects of spasticity providing insight into the many paradigms related to its pathophysiology.

REFERENCES

1. Morrell DS, Pearson JM, Sauser DD. Progressive Bone and Joint Abnormalities of the Spine and Lower Extremities in Cerebral Palsy. Radiographics 2002; 22: 257–68.

2. Malhotra S, Pandyan A, Day CR, Jones PW, Hermens H. Spasticity, an impairment that is poorly defined and poorly measured. Clin Rehabil 2009; 23: 651–8.

3. Tardieu G, Shentoub S, Delarue R. A la recherche d'une technique de mesure de la spasticite imprime avec le periodique. Neurologique 1954; 91: 143–4.

4. Lance JW. The control of muscle tone, reflexes, and movement: Robert Wartenbeg Lecture. Neurology 1980; 30: 1303.

5. Pandyan A, Gregoric M, Barnes M, Wood D, Wijck F Van, Burridge J, et al. Spasticity: Clinical perceptions, neurological realities and meaningful measurement. Disabil Rehabil 2005; 27: 2–6.

6. Sanger TD, Delgado MR, Gaebler-Spira D, Hallett M, Mink JW. Classification and definition of disorders causing hypertonia in childhood. Pediatrics 2003; 111: e89–e97.

7. Fleuren JFM, Voerman GE, Erren-Wolters C V, Snoek GJ, Rietman JS, Hermens HJ, et al. Stop using the Ashworth Scale for the assessment of spasticity. J Neurol Neurosurg Psychiatry 2010; 81: 46–52.

8. Bohannon RW, Smith MB. Interrater reliability of a modified Ashworth scale of muscle spasticity. Phys Ther 1987; 67: 206–7.

9. Boyd RN, Graham HK. Objective measurement of clinical findings in the use of botulinum toxin type A for the management of children with cerebral palsy. Eur J Neurol 1999; 6: 23–35.

10. Haugh AB, Pandyan AD, Johnson GR. A systematic review of the Tardieu Scale for the measurement of spasticity. Disabil Rehabil 2006; 28: 899–907.

11. Platz T, Eickhof C, Nuyens G, Vuadens P. Clinical scales for the assessment of spasticity, associated phenomena, and function: a systematic review of the literature. Disabil Rehabil 2005; 27: 7–18.

12. Van den Noort JC, Scholtes VA, Harlaar J. Evaluation of clinical spasticity assessment in cerebral palsy using inertial sensors. Gait Posture 2009; 30: 138–43.

13. Pandyan AD, Van Wijck FMJ, Stark S, Vuadens P, Johnson GR, Barnes MP. The construct validity of a spasticity measurement device for clinical practice: an alternative to the Ashworth scales. Disabil Rehabil 2006; 28: 579–85.

14. Lebiedowska MK, Fisk JR. Passive dynamics of the knee joint in healthy children and children affected by spastic paresis. Clin Biomech 1999; 14: 653–60.

15. Bar-On L, Aertbeliën E, Molenaers G, Bruyninckx H, Monari D, Jaspers E, et al. Comprehensive quantification of the spastic catch in children with cerebral palsy. Res Dev Disabil 2012; 34: 386–96.

16. Shindler-Ivens S, Shields R. Soleus H-reflex recruitment is not altered in persons with chronic spinal cord injury. Arch Phys Med Rehabil 2004; 85: 840–7.

17. Biering-Sørensen F, Nielsen JB, Klinge K. Spasticity-assessment: a review. Spinal Cord 2006; 44: 708–22.

18. Wood D, Burridge J, Van Wijck F, McFadden C, Hitchcock R, Pandyan A, et al. Biomechanical approaches applied to the lower and upper limb for the measurement of spasticity: A systematic review of the literature. Disabil Rehabil 2005; 27: 19–33.

19. Voerman G, Gregorič M, Hermens H. Neurophysiological methods for the assessment of spasticity: The Hoffmann reflex, the tendon reflex, and the stretch reflex. Disabil Rehabil 2005; 27: 33–68.

20. Burridge J, Wood D, Hermens H, Voerman G, Johnson G, Van Wijck F, et al. Theoretical and methodological considerations in the measurement of spasticity. Disabil Rehabil 2005; 27: 69–80.

21. Sinkjaer T, Magnussen I. Passive, intrinsic and reflex-mediated stiffness in the ankle extensors of hemiparetic patients. Brain 1994; 117: 355–63.

22. Mirbagheri MM, Barbeau H, Ladouceur M, Kearney RE. Intrinsic and reflex stiffness in normal and spastic, spinal cord injured subjects. Exp Brain Res 2001; 141: 446–59.

23. Chung SG, van Rey E, Bai Z, Rymer WZ, Roth EJ, Zhang L-Q. Separate quantification of reflex and nonreflex components of spastic hypertonia in chronic hemiparesis. Arch Phys Med Rehabil 2008; 89: 700–10.

24. De Vlugt E, de Groot JH, Schenkeveld KE, Arendzen JH, van der Helm FCT, Meskers CGM. The relation between neuromechanical parameters and Ashworth score in stroke patients. J Neuroeng Rehabil 2010; 7: 35.

25. Rabita G, Dupont L, Thevenon A, Lensel-Corbeil G, Pérot C, Vanvelcenaher J. Differences in kinematic parameters and plantarflexor reflex responses between manual (Ashworth) and isokinetic mobilisations in spasticity assessment. Clin Neurophysiol 2005; 116: 93–100.

26. Lee H-M, Chen J-JJ, Ju M-S, Lin C-CK, Poon PPW. Validation of portable muscle tone measurement device for quantifying velocity-dependent properties in elbow spasticity. J Electromyogr Kinesiol 2004; 14: 577–89.

27. Wu Y-N, Ren Y, Goldsmith A, Gaebler D, Liu SQ, Zhang L-Q. Characterization of spasticity in cerebral palsy: dependence of catch angle on velocity. Dev Med Child Neurol 2010; 52: 563–9.

28. Bénard MR, Jaspers RT, Huijing PA, Becher JG, Harlaar J. Reproducibility of hand-held ankle dynamometry to measure altered ankle moment-angle characteristics in children with spastic cerebral palsy. Clin Biomech; 2010; 25: 802–8.

29. Flamand VH, Massé-Alarie H, Schneider C. Psychometric evidence of spasticity measurement tools in cerebral palsy children and adolescents: a systematic review. J Rehabil Med 2013; 45: 14–23.

30. Johnson GR. Outcome measures of spasticity. Eur J Neurol 2002; 9 Suppl 1: 10–6.

31. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, et al. The COSMIN checklist for assessing the methodological quality of studies on

measurement properties of health status measurement instruments: an international Delphi study. Qual Life Res 2010; 19: 539–49.

32. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. J Clin Epidemiol; 2010; 63: 737–45.

33. Terwee CB, Mokkink LB, Knol DL, Ostelo RWJG, Bouter LM, de Vet HCW. Rating the methodological quality in systematic reviews of studies on measurement properties: a scoring system for the COSMIN checklist. Qual Life Res 2012; 21: 651–7.

34. Terwee CB, Bot SDM, de Boer MR, van der Windt DAWM, Knol DL, Dekker J, et al. Quality criteria were proposed for measurement properties of health status questionnaires. J Clin Epidemiol 2007; 60: 34–42.

35. Voerman GE, Burridge JH, Hitchcock R a, Hermens HJ. Clinometric properties of a clinical spasticity measurement tool. Disabil Rehabil 2007; 29: 1870–80.

36. Pandyan AD, Vuadens P, van Wijck FM, Stark S, Johnson GR, Barnes MP. Are we underestimating the clinical efficacy of botulinum toxin (type A)? Quantifying changes in spasticity, strength and upper limb function after injections of Botox to the elbow flexors in a unilateral stroke population. Clin Rehabil 2002; 16: 654–60.

37. Malhotra S, Cousins E, Ward A, Day C, Jones P, Roffe C, et al. An investigation into the agreement between clinical, biomechanical and neurophysiological measures of spasticity. Clin Rehabil 2008; 22: 1105–15.

38. Chen J-JJ, Wu Y-N, Huang S-C, Lee H-M, Wang Y-L. The use of a portable muscle tone measurement device to measure the effects of botulinum toxin type a on elbow flexor spasticity. Arch Phys Med Rehabil 2005; 86: 1655–60.

39. Turk R, Notley SV, Pickering RM, Simpson DM, Wright PA, Burridge JH. Reliability and sensitivity of a wrist rig to measure motor control and spasticity in poststroke hemiplegia. Neurorehabil Neural Repair 2008; 22: 684–96.

40. Ada L, Vattanasilp W, O'Dwyer NJ, Crosbie J. Does spasticity contribute to walking dysfunction after stroke? J Neurol Neurosurg Psychiatry 1998; 64: 628–35.

41. Vattanasilp W, Ada L, Crosbie J. Contribution of thixotropy, spasticity, and contracture to ankle stiffness after stroke. J Neurol Neurosurg Psychiatry 2000; 69: 34–9.

42. Van der Salm A, Veltink PH, Hermens HJ, Ijzerman MJ, Nene AV. Development of a new method for objective assessment of spasticity using full range passive movements. Arch Phys Med Rehabil 2005; 86: 1991–7.

43. Lamontagne A, Malouin F, Richards CL, Dumas F. Evaluation of reflex- and nonreflex-induced muscle resistance to stretch in adults with spinal cord injury using hand-held and isokinetic dynamometry. Phys Ther 1998; 78: 964–78.

44. Alhusaini AA, Dean CM, Crosbie J, Shepherd RB, Lewis J. Evaluation of spasticity in children with cerebral palsy using Ashworth and Tardieu Scales compared with laboratory measures. J Child Neurol 2010; 25: 1242–7.

45. Bar-On L, Aertbeliën E, Wambacq H, Severijns D, Lambrecht K, Dan B, et al. A clinical measurement to quantify spasticity in children with cerebral palsy by integration of multidimensional signals. Gait Posture 2012; 38: 141–7.

46. Bar-On L, Van Campenhout A, Desloovere K, Aertbeliën E, Huenaerts C, Vandendoorent B, et al. Is an instrumented spasticity assessment an improvement over clinical spasticity scales in assessing and predicting the response to integrated Botulinum Toxin-A treatment in children with Cerebral Palsy? Arch Phys Med Rehabil; 2013; doi: 10.10.

47. Lebiedowska MK, Fisk JR. Knee resistance during passive stretch in patients with hypertonia. J Neurosci Methods 2009; 179: 323–30.

48. Thilmann AF, Fellows SJ, Garms E. The mechanism of spastic muscle hypertonus. Variation in reflex gain over the time course of spasticity. Brain 1991; 114: 233–44.

49. Nielsen JB, Petersen NT, Crone C, Sinkjaer T. Stretch reflex regulation in healthy subjects and patients with spasticity. Neuromodulation 2005; 8: 49–57.

50. Dietz V, Sinkjaer T. Spastic movement disorder: impaired reflex function and altered muscle mechanics. Lancet Neurol 2007; 6: 725–33.

51. Phadke CP, Ismail F, Boulias C. Assessing the neurophysiological effects of botulinum toxin treatment for adults with focal limb spasticity: a systematic review. Disabil Rehabil 2012; 34: 91–100.

52. Willerslev-Olsen M, Lorentzen, Nielsen JB, Sinkjaer T. Passive muscle properties are altered in children with cerebral palsy before the age of 3 years and are difficult to distinguish clinically from spasticity. Dev Med Child Neurol 2013; 55: 617–23.

53. De Vet HC, Terwee CB, Ostelo RW, Beckerman H, Knol DL, Bouter LM. Minimal changes in health status questionnaires: distinction between minimally detectable change and minimally important change. Health Qual Life Outcomes 2006; 4: 54.

54. Delp SL, Anderson FC, Arnold AS, Loan P, Habib A, John CT, et al. OpenSim: open-source software to create and analyze dynamic simulations of movement. IEEE Trans Biomed Eng 2007; 54: 1940–50.

55. Levin MF, Feldman AG. The role of stretch reflex threshold regulation in normal and impaired motor control. Brain Res 1994; 657: 23–30.

56. Musampa NK, Mathieu PA, Levin MF. Relationship between stretch reflex thresholds and voluntary arm muscle activation in patients with spasticity. Exp brain Res 2007; 181: 579–93.

57. Jobin A, Levin MF. Regulation of stretch reflex threshold in elbow flexors in children with cerebral palsy: a new measure of spasticity. Dev Med Child Neurol 2000; 42: 531–40.

58. Calota A, Feldman AG, Levin MF. Spasticity measurement based on tonic stretch reflex threshold in stroke using a portable device. Clin Neurophysiol 2008; 119: 2329–37.

59. Alibiglou L, Rymer WZ, Harvey RL, Mirbagheri MM. The relation between Ashworth scores and neuromechanical measurements of spasticity following stroke. J Neuroeng Rehabil 2008; 5: 18.

60. Yam WKL, Leung MSM. Interrater Reliability of Modified Ashworth Scale and Modified Tardieu Scale in Children With Spastic Cerebral Palsy. J Child Neurol 2006; 21: 1031–5.

61. Pandyan AD, Johnson GR, Price CIM, Curless RH, Barnes MP, Rodgers H. A review of the properties and limitations of the Ashworth and modified Ashworth Scales as measures of spasticity. Clin Rehabil 1999; 13: 373–83.

62. Condliffe EG, Clark DJ, Patten C. Reliability of elbow stretch reflex assessment in chronic post-stroke hemiparesis. Clin Neurophysiol 2005; 116: 1870–8.

63. Staude G, Wolf W. Objective motor response onset detection in surface myoelectric signals. Med Eng Phys 1999; 21: 449–67.

64. Pandyan AD, Price CI, Rodgers H, Barnes MP, Johnson GR. Biomechanical examination of a commonly used measure of spasticity. Clin Biomech 2001; 16: 859–65.

Figure Legend

Figure 1. Flow chart of article search and selection strategy. *References refer to the reference list in SuppInfo4.

Table 1 Characteristics of included studies: study populations and protocol design

	Study population						Protocol design							
First author	Subjects	N	Age in years	Diagnostic details	Functional level	Main Selection Criteria	Agonists tested for spasticity	Antagonists	Instruments	Type and trajectory of stretch	Stretch velocities	Number of reps per velocity	Rest period between reps	Comparator tests
Lamontagne 1998[43]	SCI	9	Mean 41 SD 11	1-5 years post injury; complete (n=8); incomplete (n=2); traumatic (n=8); ischemic (n=1)	C6 (n=1); T5-T6 (n=1); T5 (n=3); T7 (n=1); T8 (n=1); T10 (n=2)	MAS score ≥1; no fixed contractures or deformities in lower limbs; no history of fracture or thrombophlebitis	Soleus	Tibialis anterior	Hand-held dynamometer; electrogoniometer and potentiometer; sEMG; metronome	Ramp movement from -35° plantarflexion to 5° dorsiflexion	Low velocity average: 3.3 SD 3.4°/s; high velocity average: 311.1 SD 380°/s	5	1 sec	Kin-Kom isokinetic dynamometer
Wu 2010[27]	CP	10	Mean: 10 SD 3	1 quadriplegia; 6 RH; 3 LH Movement disorder (spasticity, dystonia, ataxia) not mentioned	GMFCS: I (n=2); II (n=3); III (n=2); IV (n=2); V (n=1) MACS: II (n=5); III (n=4); V (n=1)	Not mentioned	Bicpes brachii	Tricpes brachii	Torque sensor, potentiometer, sEMG	Ramp movement from full elbow flexion to full elbow extension	30, 90,180, 270°/s	1 at 30°/s, 3 at 90°/s, 180°/s, and 270°/s	1 min	MAS
	TD	10	Mean: 10 SD 3	NR	NR									
Voerman 2007[35]	Stroke	12	Mean: 57 SD 9	First stroke, 9 LH; 3 RH	ARAT: (scored for 6 subjects) 0 (n=3); 2 (n=1); 5 (n=1); 6 (n=1)	AS 1-3 in wrist and finger flexors, >20º pain-free wrist extension, 5º active wrist flexion, able to communicate, no history of serious medical, psychological or cognitive impairment	Wrist flexors	Wrist extensors	Hand-held dynamometer, potentiometer, sEMG, electronic metronome	Sinusoidal wrist movement from neutral to extension and back to neutral	30, 60, 90 cycles/min (180, 360, 540°/s)	5-7	None	MAS; ARAT; wrist rig
	Healthy subjects	11	Mean: 57 SD 8	NR	NR	Not mentioned								

	Study population						Protocol design							
First author	Subjects	N	Age in years	Diagnostic details	Functional level	Main Selection Criteria	Agonists tested for spasticity	Antagonists	Instruments	Type and trajectory of stretch	Stretch velocities	Number of reps per velocity	Rest period between reps	Comparator tests
Van der Salm 2005[42]	SCI	9	Mean: 35 SD 7	Minimum 6 months after injury	C5 (n=1); C6 (n=2); C6-7 (n=1); T4 (n=1); T5 (n=1); T4-5 (n=1); T8 (n=1); T11 (n=1)	MAS ≥1, >18 years, absence of voluntary movements in triceps surae, tibialis anterior can contract using electrical stimulation, no fixed ankle contracture	Triceps surae	none	Calibrated strain gauge dynanometer, potentiometer, gyroscope, sEMG	Ramp movement across full ankle ROM	Random between 30-150°/s	30-40	5 sec	MAS
Bar-On 2012[45]	CP	28	Mean 10 SD 5	Spastic CP; 3 RH; 5 LH; 19 diplegia; 1 quadriplegia	GMFCS: I (n=10); II (n=12); III (n=5); IV (n=1)	Age 5-18; spastic CP; no ankle or knee contractures, no previous orthopedic surgery, no intrathecal baclofen pump; no SDR; no BTX in last 6 months	Medial gastrocnemius; medial hamstrings	Tibialis anterior; rectus femoris	Torque/force load-cell; inertial measurement units, sEMG	Ramp movement across full ankle or knee ROM	Average low velocity: Gas. 22.5 SD 7.2°/s; Hams. 35.2 SD 7.5°/s; Average high velocity: Gas. 202.1 SD 54.2°/s; Hams. 317.7 SD 47.7°/s	4	7 sec	MAS
	TD	10	Mean 11 SD 6	NR	NR	Not mentioned								
Bar-On 2013[46]	CP	31	Mean 9 SD 2	Spastic CP; 6 RH; 5 LH; 17 diplegia, 1 triplegia; 2 quadriplegia	GMFCS: I (n=12); II (n=12); III (n=6); IV (n=1)	Age 3-18; spastic CP; no ankle or knee contractures, no previous orthopedic surgery, no intrathecal baclofen pump; no SDR	Medial hamstrings	Rectus femoris	Torque/force load-cell; inertial measurement units, sEMG	Ramp movement across full knee ROM	Average low velocity: 75.48 SD 17.31°/s; average high velocity: 288.44 SD 54.11°/s	4	7 sec	MAS; MTS
Pandyan 2006[13]	Stroke	14	Median: 61 IQR 52-63	Median 48 months post stroke (IQR 32-60), 6 LH; 8 RH	Not mentioned	Clinical diagnosis of spasticity, capable of providing written, informed consent	Bicpes brachii	Triceps brachii	Force tranducer, electrogonio-meter, sEMG	Ramp movement across full elbow ROM with humerus abducted to 90°	Slow, fast (median difference: 34˚/s IQR 20-46˚/s)	1 slow stretch, 1 fast stretch	Not mentioned	MAS

	Study population						Protocol design							
First author	Subjects	N	Age in years	Diagnostic details	Functional level	Main Selection Criteria	Agonists tested for spasticity	Antagonists	Instruments	Type and trajectory of stretch	Stretch velocities	Number of reps per velocity	Rest period between reps	Comparator tests
Lebiedowska 2009[47]	Stroke	3	Mean: 65 SD 8	Not mentioned	Not mentioned	Not mentioned	Medial hamstrings, Rectus femoris	Rectus femoris, Medial hamstrings	Hand-held stain gauge dynanonmeter, potentiometer, sEMG	Ramp movement from neutral to knee extension and from neutral to 142˚ knee flexion	0.2-1.5 rad/s (11.5-540˚/s)	Several (not reported in detail)	Not mentioned	none
	Adults with CP	4	35 SD 12	3 diplegia; 1 RH										
	Children with CP	13	13 SD 4	10 diplegia; 2 RH; 1 LH										
	Healthy subjects	19	13 SD 8											
Fleuren 2010[7]	Stroke	18	Mean: 57 SD 13-16	Not mentioned	Not mentioned	Self-reported spasticity, no contractures, no severe pain, able to understand simple commands	Biceps brachii, Brachio-radialis, Rectus femoris, Vastus lateralis	none	Hand-held dynanometer, electrogonio-meter, sEMG	Ramp movement across full elbow and knee ROM (patient sidelying)	Slow, fast (median velocity: 76.6˚/s for elbow flexors, 85.2˚/s for knee extensors)	1 at slow velocity, 2 at fast velocity	Not mentioned	AS
	CP	1												
	SCI	2												
	NMD	4												
Malhotra 2008[37]	Stroke	100	Median 74 IQR 43-91	Average of 3 weeks post first stroke (range 1-6) 52 RH; 48 LH	ARAT: 0 (n=97); 1 (n=2); 3 (n=1)	Within 6 weeks of first stroke, score of 0 on grasp section of ARAT, no wrist contractures, no major illness	Long wrist flexors	Long wrist extensors	Force tranducer, electrogonio-meter, sEMG	Ramp movement across full wrist ROM	Slow, fast (mean difference between velocities: 87˚/s, SD 36˚/s, range 10-190˚/s)	1 at each velocity	NR	MAS, ARAT, BI
Chen 2005[38]	Stroke	10	Mean: 57 SD 12	Average if 38±27 months post stroke, 3 RH; 7 LH	BI: III (n=4); IV (n=2); V (n=4)	At least 6 months post stroke, no elbow contractures, no severe cognitive or affective dysfunction, BI≥ III	Biceps brachii	Triceps brachii	Air bags, differential pressure sensor, angular velocity sensor, sEMG	Sinusoidal movement from 120˚ to 60˚ elbow flexion	1/3, 1/2, 1, 1.5 Hz (120, 180, 360, 540˚/s)	Not mentioned	≥30 sec	MAS
Turk 2008[39]	Stroke	12	Mean: 62 SD 12	6±4 years post stroke, 4 RH; 8 LH	Mean ARAT: 18.8±11.5	At least 3 months post stroke, some active wrist movement, no wrist contractures, no neglect or major illness	Flexor carpi ulnaris, Flexor carpi radialis	Extensor carpi radialis longus	Strain gauges (force sensor), potentiometer, sEMG	Sinusoidal movement across full wrist ROM	Slow: 0.04 or 0.08Hz (14.4 or 28.8˚/s) Fast: 1.5Hz (540˚/s)	2 at slow velocity followed by fast sinusoidal	Not mentioned	MAS, ARAT
	Healthy adults	12	51 SD 20	NR	NR	Not mentioned								

	Study population						Protocol design							
First author	Subjects	N	Age in years	Diagnostic details	Functional level	Main Selection Criteria	Agonists tested for spasticity	Antagonists	Instruments	Type and trajectory of stretch	Stretch velocities	Number of reps per velocity	Rest period between reps	Comparator tests
Alhusaini 2010[44]	CP	27	Mean: 7 SD 2	Not mentioned	GMFCS I and II	Spastic CP, GMFCS I-II, no severe cognitive dysfunction, no orthopedic surgery, or anti-spasticity treatment in previous 5 months	Medial gastrocnemius, Soleus	Tiabialis anterior	Load cell, potentiometer, sEMG	Ramp movement across full ROM	As slow as possible, Slow, Fast	At least 3	Not mentioned	MAS, TS
Ada 1998[40]	Stroke	14	Mean 65 SD 9	Hemiparetic; 5-10 months post stroke	≥3 on motor assessment scale	≥3 on motor assessment scale; sufficient cognitive ability	Medial gastrocnemius	None	Load cell, potentiometer, sEMG	Sinusoidal between 10° plantarflexion and 10° dorsiflexion	0.5, 1, 1.5, 2Hz (180, 360, 540, 720°/s)	Each velocity trial was performed during 25 sec	None	None
	Healthy	15	Mean 52 SD 6	NR	NR	Neurologically normal								
Vattanaslip 2012[41]	Stroke	30	Mean 68 SD 9	2-5 years post stroke, 12 RH, 18 LH	Not mentioned	Calf muscles diagnosed as clinically stiff, ≥2 AS, sufficient cognitive ability, no other problems interfering with ankle motion	Medial gastrocnemius	none	Load cell, potentiometer, sEMG	Ramp movement across full ROM; Sinusoidal between 10° plantarflexion and 10° dorsiflexion	Undefined velocity for assessing contracture, 2°/s for assessing thixotropy, and at 2Hz (720°/s) for assessing spasticity	1 at undefined velocity; 2 at 2°/s and during 30 sec at 2Hz (720°/s)	None	AS
	Healthy subjects	10	Mean 59 SD 8	NR	NR	Not mentioned								

Reps, repetitions; CP, cerebral palsy; TD, typically developing; SCI, spinal cord injury; NMD, neuromuscular disease; RH, right hemiplegia; LH, left hemiplegia; IQR, Inter quartile range; reps., repetitions; GMFCS, Gross Motor Function Classification System; MACS, Manual Ability Classification System; ARAT, Action Research Arm Test; (M)AS, (Modified) Ashworth Score; (M)TS, (Modified) Tardieu score; BI, Brunnstrom Index; ROM, range of motion; sEMG, surface electromyography

Table 2 Outcome parameters from instrumented tests developed from different signals at different stretch velocities

First author	Position	Torque	sEMG
Lamontagne 1998[43]	Low velocity Average angular velocity at -5° plantarflexion High velocity Maximum angular velocity	Low and high velocity Average torque at -5° plantarflexion	High velocity EMG onset was defined when EMG > 2SD than mean baseline level preceding onset
Wu 2010[27]	30˚/sec ROM AOC = angle at maximum (dr(T)/dt) Ratio between AOC and ROM	30˚/sec Slope of torque-angle curve at 70° elbow flexion Energy loss: area between ascending and descending limbs of torque-angle curve Torque at 45°, 60°, 75° elbow flexion At 90,180, 270°/s Slope of peak torque vs. 3 stretch velocities Peak torque Maximum (dr(T)/dt)	90,180, 270°/s EMG onset angle
Voerman 2007[35]	Slow Passive wrist ROM 30, 60, 90cycles/min (180, 360, 540°/s) Passive wrist extension ROM Angular velocity	30, 60, 90cycles/min (180, 360, 540°/s) Slope of torque-angle curve from neutral to full wrist extension	30, 60, 90 cycles/min (180, 360, 540°/s) Average RMS-EMG from neutral to full wrist extension
Van der Salm 2005[42]	<70°/s ROM High velocity ROM Average maximum angular velocity	<70°/s Average torque in 3 zones over the full ROM	50, 75, 100 °/s Average RMS-EMG over 100ms window after EMG onset (>3SD) plotted against stretch velocities, exponential fit over 30-45 values Angle and angular velocity at EMG onset Slope values of angle/velocity onsets Angle at 100°/s = reflex initiating angle
Bar-On 2012[45]	Low velocity ROM Average maximum angular velocity	Low and high velocity Change in average torque at maximum velocity between velocity trials Change in average integral of torque-angle curve from max. velocity to 90% ROM between velocities	Peak of three MVICs Low and high velocity Change in in average RMS-EMG in maximum velocity zone (200ms before max. velocity to 90% ROM) between velocity trials (expressed as % of peak value of three maximum voluntary isometric contractions) EMG onset defined as time of first muscle activity according to method of Staude & Wolf[63]
Bar-On 2013	Low velocity ROM High velocity Average maximum angular velocity AOC defined as the angle corresponding to the time of minimum power after maximum power during first high velocity stretch, expressed as % of ROM	Low and high velocity Change in average torque at 70° knee flexion between velocity trials Change in average integral of torque-angle curve from max. velocity to 90% ROM between velocity trials High velocity Minimum power after maximum power in first high velocity stretch	Low and high velocity Change in in average RMS-EMG in maximum velocity zone (200ms before max. velocity to 90% ROM) between velocity trials
Pandyan 2006[13]	Slow and fast ROM Average angular velocity	Slow and fast Change in slope of force-angle curve between velocities over full ROM	Slow and fast Change in RMS-EMG over full ROM between velocities

First author	Position	Torque	sEMG
Lebiedowska 2009[47]	0.2-1.5 rad/s (11.5˚/s - 540˚/s) Passive ROM	0.2-1.5 rad/s (11.5-540˚/s) Slope of torque-angle curve during initial increase Integral of torque-angle curve over full ROM	0.2-1.5 rad/s (11.5-540˚/s) Maximum value of RMS-EMG over ROM Slope of RMS-EMG velocity curve Hypertonia of neural origin: RMS-EMG ≥ mean ± 3SD before movement began in slow and fast velocity stretches. Hypertonia of non-neural origin: RMS-EMG < mean ±3 SD before movement began in slow and fast velocity stretches.
Fleuren 2010[7]	Slow Passive ROM	Slow and fast Integral of torque-time curve over full ROM	Slow and fast Average RMS-EMG over full ROM
Malhotra 2008[37]	Slow and fast ROM	Slow and fast Slope of force-angle curve 10-90% ROM Shapes of force-angle curves: • Slope of force-angle curve <0.7N/˚: neg. stiffness. • Slope of force-angle curve > 0.7N/˚ and R^2>0.6: linear stiffness • Slope of force-angle curve >0.7N/˚and R^2<0.6: catch or clasp-knife): non-linear stiffness	Slow and fast Average RMS-EMG over full ROM Patterns of muscle response: • No/negligible muscle response • Position-dependent: muscle response independent of stretch velocity • Velocity-dependent: negligible muscle activity during slow stretch, increased activity during fast stretch • Position- and velocity-dependent • Early catch: early muscle activation reducing as the muscle lengthens
Chen 2005[38]	1/3, 1/2, 1, 1.5Hz (120˚/s, 180˚/s, 360˚/s, 540˚/s) ROM	1/3, 1/2, 1, 1.5Hz (120˚/s, 180˚/s, 360˚/s, 540˚/s) Velocity-dependent viscous component of torque (see appendix Chen 2005) Slope of viscosity-velocity graph (see Chen 2004)	1, 1.5Hz (360˚/s, 540˚/s) Angle at EMG onset
Turk 2008[39]	0.5Hz (28.6˚/s) Tracking index: ability to accurately follow tracking signal ROM	0.04Hz (14.4˚/s). Force/torque angle index: average change in force/torque between 0 and 30° wrist extension	1.5Hz (540˚/s) Stretch Index: average RMS-EMG minus resting EMG during wrist extension
Alhusaini 2010[44]	Slow ROM	Slow Contracture: angle <10° dorsiflexion at 4.6Nm of force	Fast Average, normalized RMS-EMG
Ada 1998[40]	None	0.5, 1, 1.5, 2Hz (180, 360, 540, 720°/s) Change in torque over 20° interval	0.5, 1, 1.5, 2Hz (180, 360, 540, 720°/s) Gain in RMS-EMG over ROM (µV/°)
Vattanaslip 2000[41]	Undefined low velocity ROM	2Hz (720°/s) Change in torque over 20° interval	2Hz (720°/s) Gain in RMS-EMG over ROM (µV/°)

ROM, range of motion; AOC, angle of catch; dr(T)/dt, change in torque over change in time; sEMG, surface electromyography; RMS-EMG, root mean square electromyography; MVIC, maximum isometric voluntary contraction; neg., negligible

Table 3. COSMIN scores and reasoning for scores on the reliability of included studies (for an extended version including statistical findings, see SuppInfo2)

First author	Inter-rater reliability	Intra-rater reliability	Measurement error
Lamontagne 1998[43]	Not performed	Within one session, 1sec between repetitions	Within one session, 1 sec between repetitions
COSMIN score	NA	Poor	Fair
		- only biomechanical parameter assessed for reliability; - short time interval between repetitions	- the absolute measurement error was not provided; - only biomechanical parameter assessed for reliability
Generalizability	NA	Good	Good
		- no information on missing values	- no information on missing values
Wu 2010[27]	Not performed	1 day between measurements	Not calculated
COSMIN score	NA	Fair	NA
		- only biomechanical parameter assessed for reliability	
Generalizability	NA	Poor	NA
		- reliability was only measured in typically developing children (the results cannot be generalized to a patient population)	
Voerman 2007[35]	1 day between measurements	10 minutes between measurements	Not calculated
COSMIN score	Fair	Good	NA
	- small sample	- unclear whether administrations were independent	
Generalizability	Good	Excellent	NA
	- subjects were missing an ARAT score		
Van der Salm 2005[42]	Not performed	Within one session, 5 seconds rest between repetitions	Not calculated
COSMIN score	NA	Poor	NA
		- short time interval between repetitions; - only one parameter was assessed for reliability	
Generalizability	NA	Excellent	NA
Bar-On 2012[45]	Not performed	Average of 13 SD 9 days between measurements	Average of 13 SD 9 days between measurements
COSMIN score	NA	Good	Good
		- small sample size; - no indication if subjects were stable in interim period	– small sample size. - no indication if subjects were stable in interim period. - MIC not reported
Generalizability	NA	Excellent	Excellent
Bar-On 2013[46]	Not performed **(0)**	Average of 13 SD 9 days between measurements	Average of 13 SD 9 days between measurements
COSMIN score	NA	Good	Good
		- no indication if subjects were stable in interim period	- no indication if subjects were stable in interim period . - MIC not reported
Generalizability	NA	Excellent	Excellent
Turk 2008[39]	Immediately following assessment by first rater	Interval of one measurement procedure (time not specified)	Intra-rater stroke:
COSMIN score	Good	Good	Good
	- no ICC values calculated	- time interval between administrations unknown	- for some parameters average difference between persons with disabilities and controls >SDC. - MIC not reported
Generalizability	Excellent	Excellent	Excellent
Pandyan 2001[64]	Not performed **(0)**	Within one session, 10-15 sec between repetitions	Not calculated **(0)**
COSMIN score	NA	Poor	NA
		- only biomechanical parameter assessed for reliability; - short time interval between repetitions; - no ICCs calculated	
Generalizability	NA	Excellent	NA

NA, Not Applicable; SDC, Smallest Detectable Change; MIC. Minimally Important Change; ICC, Intra Correlation Coefficient

Table 4. COSMIN scores and reasoning for scores on the validity of included studies (for an extended version including results and statistical findings, see SuppInfo 3)

First author	Content Validity	Construct validity/hypothesis testing	Responsiveness	Interpretability
Lamontagne 1998[43]	Not measured	Comparison to motor-controlled device	Not measured	Means and standard deviations of outcome parameters provided
COSMIN score	NA	Fair - small sample; - high velocity stretches not comparable between hand-held dynamometer and motor-controlled device; – description of the parameters of a motor-controlled device missing	NA	Good - small sample; -SDC and MIC not reported; – limited focus
Generalizability	NA	Good - small sample	NA	Good - small sample
Wu 2010[27]	Relation between signals Relation of signals to velocity	Comparison to control group Comparison to clinical scales	Not measured	Means and standard deviations of outcome parameters provided
COSMIN score	Good - type of cerebral palsy (spastic, dystonia, etc) not mentioned, - no description of missing data	Good - parametric statistics performed to compare groups while sample size was relatively small and data distribution not reported	NA	Fair - no description of missing data; SDC and MIC not reported
Generalizability	Good - type of cerebral palsy and study setting not mentioned	Good - type of cerebral palsy and study setting not mentioned	NA	Good - type of CP and study setting not mentioned
Voerman 2007[35]	Relation of signals to velocity	Comparison to control group Comparison to motor-controlled device	Not measured	Means and standard deviations or medians and ranges of outcome parameters provided
COSMIN score	Fair - theoretical framework described but statistical comparisons not performed	Good - the sample size used for the correlations with ARAT was small; - the measurement properties of the motor-controlled device/comparator instrument were not described	NA	Good - SDC and MIC not reported
Generalizability	Poor - no data on content validity available	Good - fewer subjects tested with ARAT and with the motor-controlled device	NA	Good - fewer subjects tested with ARAT and with the motor-controlled device
Van der Salm 2005[42]	Relation between signals Relation of signals to velocity	Not measured	Not measured	Means and standard deviations of outcome parameters provided
COSMIN score	Fair - torque only measured in 4 subjects and only at low velocity	NA	NA	Good - small sample; -SDC and MIC not reported
Generalizability	Good - characteristics of excluded subjects missing	NA	NA	Good - torque only measured in 4 subjects
Bar-On 2012[45]	Relation of signals to velocity	Comparison to control group Comparison to clinical scales	Not measured	Means and standard deviations of outcome parameters provided, SDC could be calculated
COSMIN score	Poor - no statistical tests performed	Good - Hypotheses not explicitly stated	NA	Good - MIC not reported
Generalizability	Fair - little data on content validity available	Excellent	NA	Excellent

First author	Content Validity	Construct validity/hypothesis testing	Responsiveness	Interpretability
Bar-On 2013[46]	NA	Comparison to clinical scales	Treatment with BTX	Means and standard deviations of outcome parameters provided, SDC provided
COSMIN score	NA	Excellent	Excellent	Good - MIC not reported
Generalizability	NA	Excellent	Excellent	Excellent
Pandyan 2006[13]	Relation between signals Relation of signals to velocity:	Comparison to clinical scales:	Not measured	Medians and ranges of outcome parameters provided
COSMIN score	Excellent	Good - no description of how missing data was handled	NA	Good - SDC and MIC not reported
Generalizability	Excellent	Excellent	NA	Excellent
Lebiedowska 2009[47]	Comparison between signals Relation of signals to velocity:	Comparison to control group	Not measured	Means and standard deviations of outcome parameters provided
COSMIN score	Fair - see comments on relation of signals to velocity in Suppinfo3; -statistical comparisons involving small samples	Fair - the excluded subjects' characteristics were not described; -EMG data was not normalized	NA	Fair - subgroup comparisons based on small samples; -SDC and MIC not reported
Generalizability	Fair - no diagnostic information, indication of spasticity severity, or functional level provided; - influence of heterogeneity between subjects not checked for	Fair - no diagnostic information, indication of spasticity severity, or functional level provided; -Influence of heterogeneity between subjects not checked for	NA	Fair - subgroup comparisons based on small samples
Fleuren 2010[7]	Relation between signals Relation of signals to velocity	Comparison to clinical scales	Not measured	Means and standard deviations of outcome parameters not provided
COSMIN score	Good - the instrumented parameters were correlated to the velocity of stretch with the intention of explaining the variability in performance rather than to test content validity; - muscle activity from antagonist muscles not measured	Good - the instrumented parameters were correlated to the AS with the intention of explaining the variability in performance rather than to test construct validity; - large influence of rater on multivariate mixed linear model with AS as dependent variable	NA	Poor - no instrumented data on spasticity presented; -SDC and MIC not reported
Generalizability	Good - disease characteristics not reported	Good - disease characteristics not reported	NA	Good - disease characteristics not reported
Malhotra 2008[37]	Relation between signals Relation of signals to velocity	Comparison to clinical scales	Not measured	Means and standard deviations of outcome parameters provided
COSMIN score	Excellent	Good - no information on missing data	NA	Good - SDC and MIC not reported
Generalizability	Excellent	Excellent	NA	Excellent
Chen 2005[38]	Relation of signals to velocity	Comparison to clinical scales	Treatment with BTX	Means and standard deviations of outcome parameters provided
COSMIN score	Poor - no statistical tests carried out	Poor - no statistical tests carried out; - no information on missing data	Poor - EMG parameter was compared pre-post on individual subject data rather than with group analysis; - some comparisons made using independent, rather than dependent group analyses	Fair - no information on missing data -No analysis of sub-groups;- important statistical flaws; -SDC and MIC not reported

First author	Content Validity	Construct validity/hypothesis testing	Responsiveness	Interpretability
Generalizability	Good - no information on missing data	NA	Good - no information on missing data	Good - no information on missing data
Turk 2008[39]	Not measured	Comparison to control group	Not measured	Means and SD deviations of outcome parameters provided. SDC can be calculated.
COSMIN score	NA	Good -The magnitude of expected differences between groups were not included in the hypotheses	NA	Excellent
Generalizability	NA	Excellent	NA	Excellent
Alhusaini 2010[44]	Relation of signals to velocity	Comparison to clinical scales	Not measured	Means and SD deviations of outcome parameters not provided.
COSMIN score	Poor - no statistical tests carried out	Good - no description regarding missing data; -The magnitude of expected correlations were not included in the hypotheses; - stretch velocities not reported	NA	Poor - no values from the instrumented test were reported; - SDC and MIC not reported
Generalizability	Excellent	Excellent	NA	Excellent
Ada 1998[40]	Relation between signals	Comparison to control group	NA	Means and standard deviations of outcome parameters provided
COSMIN score	Fair - sub-group analyses were based on small samples; - some missing statistical results	Good - some missing statistical results; - no hypotheses on expected result; - no information on how missing values were handled	NA	Good - SDC and MIC not reported; - some samples too small; - the percentage of responders who had lowest/highest possible scored not reported
Generalizability	Excellent	Excellent	NA	Excellent
Vattanaslip 2000[41]	Relation between signals	Comparison to control group Comparison to clinical scales	NA	Not all means and standard deviations of outcome parameters provided
COSMIN score	Good - spasticity not defined	Poor - low velocity stretch to evaluate ROM not defined; - gain in RMS-EMG not compared between groups; - change in torque was only assessed at high velocity; -no actual comparison to clinical scale as parameter values not compared to clinical scores;	NA	Poor - Missing some descriptive statistics related to contracture and spasticity; - SDC and MIC not reported
Generalizability	Good - Gender of included subjects not reported; - Place from which subjects were recruited not mentioned.	Good - Gender of included subjects not reported; - Place from which subjects were recruited not mentioned.	NA	Good - Gender of included subjects not reported; - Place from which subjects were recruited not mentioned.

NA, not applicable; SDC, smallest detectable change; MIC, minimally important change; ARAT, Action Research Arm Test; BTX, botulinum Toxin-A; EMG, electromyography; RMS-EMG, root mean square electromyography; AS, Ashworth Scale; ROM, range of motion